Media geométrica de matrices positivas

Jorge Antezana [a b] Eduardo Ghiglioni [a b] Demetrio Stojanoff [a b]

[a] Instituto Argentino de Matemática
CONICET
"Alberto P. Calderón"

[b] Centro de Matemática de La Plata
Universidad Nacional de La Plata

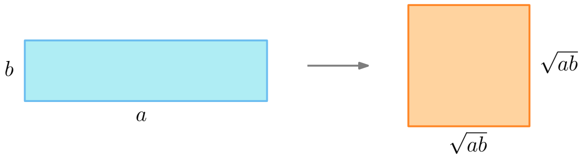
XIV Encuentro de Analistas
Alberto P. Calderón
Villa General Belgrano 2018

La media geométrica de dos números positivos a y b se define como:

$$\sqrt{ab}.$$

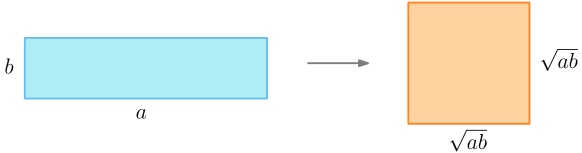
La media geométrica de dos números positivos a y b se define como:

$$\sqrt{ab}.$$



La media geométrica de dos números positivos a y b se define como:

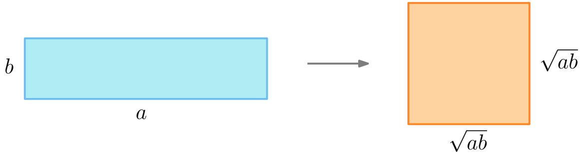
$$\sqrt{ab}.$$



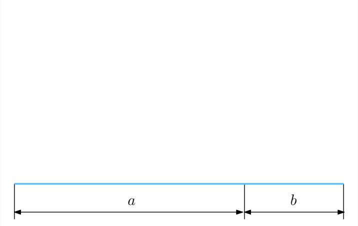
Construcción de Euclides:

La media geométrica de dos números positivos a y b se define como:

$$\sqrt{ab}.$$

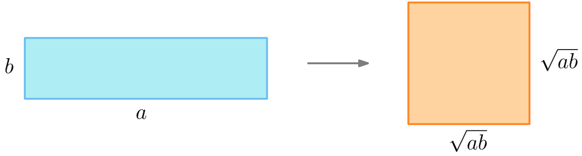


Construcción de Euclides:

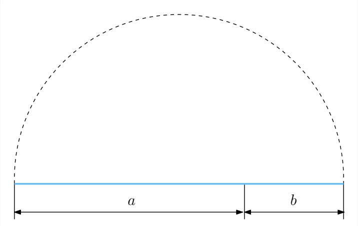


La media geométrica de dos números positivos a y b se define como:

$$\sqrt{ab}.$$

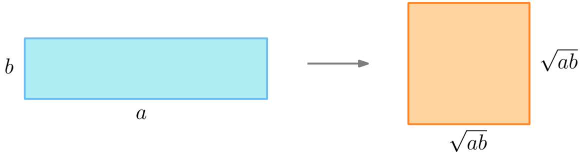


Construcción de Euclides:

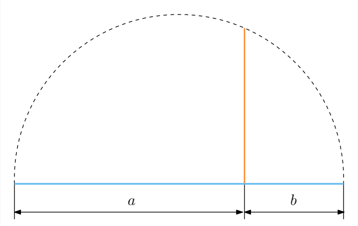


La media geométrica de dos números positivos a y b se define como:

$$\sqrt{ab}.$$

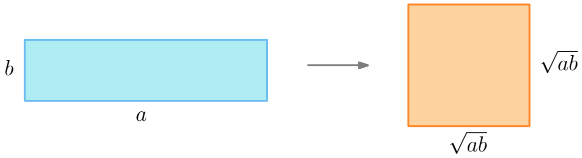


Construcción de Euclides:

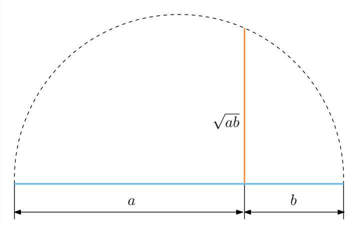


La media geométrica de dos números positivos a y b se define como:

$$\sqrt{ab}.$$

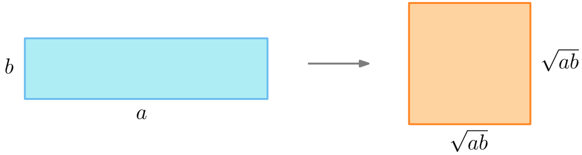


Construcción de Euclides:

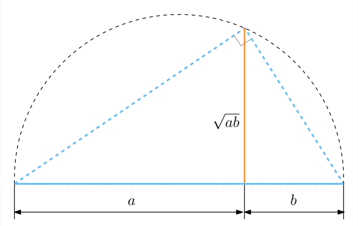


La media geométrica de dos números positivos a y b se define como:

$$\sqrt{ab}.$$



Construcción de Euclides:



Dada una matriz A de $n \times n$, decimos que es positiva si para todo $x \in \mathbb{C}^n$ no nulo

$$\langle Ax, x \rangle > 0. \tag{1.1}$$

Esto es equivalente a que $A = A^*$ y sus autovalores sean positivos. Al conjunto de matrices positivas lo denotaremos $\mathbb{P}(n)$. Diremos que A es semidefinida positiva si en (1.1) se admite la igualdad. Recordemos que dadas $A, B \in \mathcal{H}(n)$, entonces

$$A \leq B \quad \text{si} \quad B - A \geq 0.$$

Dada una matriz A de $n \times n$, decimos que es positiva si para todo $x \in \mathbb{C}^n$ no nulo

$$\langle Ax, x \rangle > 0. \tag{1.1}$$

Esto es equivalente a que $A = A^*$ y sus autovalores sean positivos. Al conjunto de matrices positivas lo denotaremos $\mathbb{P}(n)$. Diremos que A es semidefinida positiva si en (1.1) se admite la igualdad. Recordemos que dadas $A, B \in \mathcal{H}(n)$, entonces

$$A \leq B \quad \text{si} \quad B - A \geq 0.$$

Problema

Definir la media geométrica entre dos matrices positivas A y B

Dada una matriz A de $n \times n$, decimos que es positiva si para todo $x \in \mathbb{C}^n$ no nulo

$$\langle Ax, x \rangle > 0. \tag{1.1}$$

Esto es equivalente a que $A = A^*$ y sus autovalores sean positivos. Al conjunto de matrices positivas lo denotaremos $\mathbb{P}(n)$. Diremos que A es semidefinida positiva si en (1.1) se admite la igualdad. Recordemos que dadas $A, B \in \mathcal{H}(n)$, entonces

$$A \leq B \quad \text{si} \quad B - A \geq 0.$$

Problema

Definir la media geométrica entre dos matrices positivas A y B

Comentarios:

- Si $AB = BA$ entonces se puede definir la media como $\sqrt{AB}$, ya que A y B se pueden diagonalizar simultáneamente y las operaciones se realizan a nivel de los autovalores.
- Dados $A, B \in \mathbb{P}(n)$, si AB es autoadjunta entonces

$$AB = (AB)^* = B^* A^* = BA$$

Dados $a, b > 0$, entonces $\sqrt{ab}$ es el mayor número real c tal que

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} \geq 0.$$

Dados $a, b > 0$, entonces $\sqrt{ab}$ es el mayor número real c tal que

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} \geq 0.$$

Teorema (Pusz, Woronowicz '75 - Anderson, Trapp '80)

Dadas $A, B \in \mathbb{P}(n)$, el conjunto

$$\left\{ C \in \mathcal{H}(n) : \begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq 0 \right\}$$

admite un máximo. Más aún, dicho máximo es

$$A\#B = A^{1/2}(A^{-1/2}B\,A^{-1/2})^{1/2}A^{1/2} = B^{1/2}(B^{-1/2}A\,B^{-1/2})^{1/2}B^{1/2}.$$

Dados $a, b > 0$, entonces $\sqrt{ab}$ es el mayor número real c tal que

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} \geq 0.$$

Teorema (Pusz, Woronowicz '75 - Anderson, Trapp '80)

Dadas $A, B \in \mathbb{P}(n)$, el conjunto

$$\left\{ C \in \mathcal{H}(n) : \begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq 0 \right\}$$

admite un máximo. Más aún, dicho máximo es

$$A\#B = A^{1/2}(A^{-1/2}B\,A^{-1/2})^{1/2}A^{1/2} = B^{1/2}(B^{-1/2}A\,B^{-1/2})^{1/2}B^{1/2}.$$

La simetría se debe a que

$$\begin{pmatrix} B & C \\ C & A \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^* \begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Dados $a, b > 0$, entonces $\sqrt{ab}$ es el mayor número real c tal que

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} \geq 0.$$

Teorema (Pusz, Woronowicz ’75 - Anderson, Trapp ’80)

Dadas $A, B \in \mathbb{P}(n)$, el conjunto

$$\left\{ C \in \mathcal{H}(n) : \begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq 0 \right\}$$

admite un máximo. Más aún, dicho máximo es

$$A\#B = A^{1/2}(A^{-1/2}B\,A^{-1/2})^{1/2}A^{1/2} = B^{1/2}(B^{-1/2}A\,B^{-1/2})^{1/2}B^{1/2}.$$

La simetría se debe a que

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq 0 \iff \begin{pmatrix} B & C \\ C & A \end{pmatrix} \geq 0$$

Dados $a, b > 0$, entonces $\sqrt{ab}$ es el mayor número real c tal que

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} \geq 0.$$

Teorema (Pusz, Woronowicz '75 - Anderson, Trapp '80)

Dadas $A, B \in \mathbb{P}(n)$, el conjunto

$$\left\{ C \in \mathcal{H}(n) : \begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq 0 \right\}$$

admite un máximo. Más aún, dicho máximo es

$$A \# B = A^{1/2}(A^{-1/2} B\, A^{-1/2})^{1/2} A^{1/2} = B^{1/2}(B^{-1/2} A\, B^{-1/2})^{1/2} B^{1/2}.$$

La simetría se debe a que

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq 0 \iff \begin{pmatrix} B & C \\ C & A \end{pmatrix} \geq 0$$

Proposición

La función $(A, B) \mapsto A \# B$ es monótona en A y B.

Dada una matriz $G \in \mathcal{G}l\,(n)$

$$\begin{pmatrix} G^*AG & G^*CG \\ G^*CG & G^*BG \end{pmatrix} = \begin{pmatrix} G^* & 0 \\ 0 & G^* \end{pmatrix} \begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} G & 0 \\ 0 & G \end{pmatrix}$$

Dada una matriz $G \in \mathcal{G}l\,(n)$

$$\begin{pmatrix} G^*AG & G^*CG \\ G^*CG & G^*BG \end{pmatrix} = \begin{pmatrix} G^* & 0 \\ 0 & G^* \end{pmatrix} \begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} G & 0 \\ 0 & G \end{pmatrix}$$

Entonces

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq 0 \iff \begin{pmatrix} G^*AG & G^*CG \\ G^*CG & G^*BG \end{pmatrix} \geq 0$$

Dada una matriz $G \in \mathcal{G}l\,(n)$

$$\begin{pmatrix} G^*AG & G^*CG \\ G^*CG & G^*BG \end{pmatrix} = \begin{pmatrix} G^* & 0 \\ 0 & G^* \end{pmatrix} \begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} G & 0 \\ 0 & G \end{pmatrix}$$

Entonces

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq 0 \iff \begin{pmatrix} G^*AG & G^*CG \\ G^*CG & G^*BG \end{pmatrix} \geq 0$$

Proposición

Dadas $A, B \in \mathbb{P}(n)$ y $G \in \mathcal{G}l\,(n)$ entonces

$$G^*AG \# G^*BG = G^*(A \# B)G.$$

Dada una matriz $G \in \mathcal{G}l\,(n)$

$$\begin{pmatrix} G^*AG & G^*CG \\ G^*CG & G^*BG \end{pmatrix} = \begin{pmatrix} G^* & 0 \\ 0 & G^* \end{pmatrix} \begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} G & 0 \\ 0 & G \end{pmatrix}$$

Entonces

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq 0 \iff \begin{pmatrix} G^*AG & G^*CG \\ G^*CG & G^*BG \end{pmatrix} \geq 0$$

Proposición

Dadas $A, B \in \mathbb{P}(n)$ y $G \in \mathcal{G}l\,(n)$ entonces

$$G^*AG \,\#\, G^*BG = G^*(A \,\#\, B)G.$$

Observación: Con esta proposición en mente, notemos que

$$A\#B = A^{1/2}\big(\,I\#(A^{-1/2}B\,A^{-1/2})\,\big)A^{1/2} = A^{1/2}(A^{-1/2}B\,A^{-1/2})^{1/2}A^{1/2}.$$

Definición

La media geométrica de dos matrices positivas A y B de igual tamaño es la matriz

$$A\#B.$$

Definición

La media geométrica de dos matrices positivas A y B de igual tamaño es la matriz

$$A\#B.$$

Proposición

Dadas $A, B \in \mathbb{P}(n)$

$$\left(A^{-1}+B^{-1}\right)^{-1} \leq \; A\#B \; \leq \; \frac{A+B}{2}.$$

La primera desigualdad se obtiene de la segunda usando que:

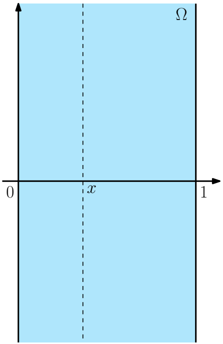
$$(A\#B)^{-1} = \left(A^{1/2}(A^{-1/2}B\,A^{-1/2})^{1/2}A^{1/2}\right)^{-1} = A^{-1}\#\,B^{-1}.$$

Sean $\|\cdot\|_0$ y $\|\cdot\|_1$ dos normas en $\mathbb{C}^n$ y Ω la banda horizontal

$$\Omega = \{z \in \mathbb{C} : 0 < \mathrm{Re}(z) < 1\}.$$

y definimos

$$\mathcal{F} = \{F : \overline{\Omega} \to \mathbb{C}^n : \ F \text{ es continua y acotada en } \overline{\Omega} \text{ y holomorfa en } \Omega\}.$$

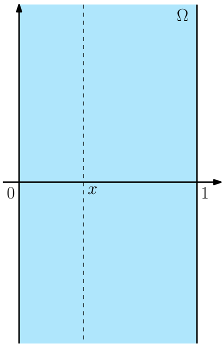


Sean $\|\cdot\|_0$ y $\|\cdot\|_1$ dos normas en $\mathbb{C}^n$ y Ω la banda horizontal

$$\Omega = \{z \in \mathbb{C} : 0 < \mathrm{Re}(z) < 1\}.$$

y definimos

$$\mathcal{F} = \{F : \overline{\Omega} \to \mathbb{C}^n : \; F \text{ es continua y acotada en } \overline{\Omega} \text{ y holomorfa en } \Omega\}.$$



Dado $x \in (0,1)$ definimos

$$\mathcal{F}_x = \{F \in \mathcal{F} : \; F(x) = 0\}.$$

A partir de este conjunto definimos las normas $\|\cdot\|_x$ del siguiente modo: dado $\xi \in \mathbb{C}^n$

$$\|\,\xi\,\|_x = \inf_{F \in \mathcal{F}_x} \sup_{y \in \mathbb{R}} \{\, \|\xi + F(x+iy)\|_0 \,,\, \|\xi + F(x+iy)\|_1 \}.$$

Si $[\cdot,\cdot]$ es un producto interno en $\mathbb{C}^n$, entonces existe $A \in \mathbb{P}(n)$ tal que

$$[\xi\,,\zeta] = \langle\, A\,\xi\,,\zeta\,\rangle =: \langle\,\xi\,,\zeta\,\rangle_A\,.$$

Si $[\cdot,\cdot]$ es un producto interno en $\mathbb{C}^n$, entonces existe $A \in \mathbb{P}(n)$ tal que

$$[\xi\,,\zeta] = \langle\, A\,\xi\,,\zeta\,\rangle =: \langle\, \xi\,,\zeta\,\rangle_A\,.$$

Supongamos ahora que

$$\begin{array}{lll} \|\cdot\|_0 & \text{está asociado a un producto interno} & [\xi\,,\zeta]_0 = \langle\, A_0\,\xi\,,\zeta\,\rangle \\ \|\cdot\|_1 & \text{está asociado a un producto interno} & [\xi\,,\zeta]_1 = \langle\, A_1\,\xi\,,\zeta\,\rangle \end{array}$$

con $A_0, A_1 \in \mathbb{P}(n)$.

Si $[\cdot,\cdot]$ es un producto interno en $\mathbb{C}^n$, entonces existe $A \in \mathbb{P}(n)$ tal que

$$[\xi\,,\zeta] = \langle\, A\,\xi\,,\zeta\,\rangle =: \langle\, \xi\,,\zeta\,\rangle_A\,.$$

Supongamos ahora que

$\|\cdot\|_0$	está asociado a un producto interno	$[\xi\,,\zeta]_0 = \langle\, A_0\,\xi\,,\zeta\,\rangle$
$\|\cdot\|_1$	está asociado a un producto interno	$[\xi\,,\zeta]_1 = \langle\, A_1\,\xi\,,\zeta\,\rangle$

con $A_0, A_1 \in \mathbb{P}(n)$.

Teorema (Semmes '88- Andruchow, Milman, Stojanoff '97)

La norma intermedia $\|\cdot\|_t$ está asociada a un producto escalar $\langle\,\cdot,\cdot\,\rangle_{A_t}$, donde

$$A_t = A_0^{1/2}(A_0^{-1/2}A_1\,A_0^{-1/2})^t A_0^{1/2}.$$

Si $[\cdot,\cdot]$ es un producto interno en $\mathbb{C}^n$, entonces existe $A \in \mathbb{P}(n)$ tal que

$$[\xi\,,\zeta] = \langle\, A\,\xi\,,\zeta\,\rangle =: \langle\,\xi\,,\zeta\,\rangle_A\,.$$

Supongamos ahora que

$\|\cdot\|_0$	está asociado a un producto interno	$[\xi\,,\zeta]_0 = \langle\, A_0\,\xi\,,\zeta\,\rangle$
$\|\cdot\|_1$	está asociado a un producto interno	$[\xi\,,\zeta]_1 = \langle\, A_1\,\xi\,,\zeta\,\rangle$

con $A_0, A_1 \in \mathbb{P}(n)$.

Teorema (Semmes '88- Andruchow, Milman, Stojanoff '97)

La norma intermedia $\|\cdot\|_t$ está asociada a un producto escalar $\langle\,\cdot,\cdot\,\rangle_{A_t}$, donde

$$A_t = A_0^{1/2}(A_0^{-1/2}A_1\,A_0^{-1/2})^t A_0^{1/2}.$$

Notación: Dado $t \in [0,1]$

$$A_0 \#_t A_1 := A_0^{1/2}(A_0^{-1/2}A_1\,A_0^{-1/2})^t A_0^{1/2}.$$

El conjunto $\mathbb{P}(n)$ es un cono convexo abierto de $\mathcal{H}(n)$.

El conjunto $\mathbb{P}(n)$ es un cono convexo abierto de $\mathcal{H}(n)$.

Proposición

Al conjunto $\mathbb{P}(n)$ se lo puede dotar de una estructura diferencial de modo que el espacio tangente en cada punto se puede identificar con $\mathcal{H}(n)$.

El conjunto $\mathbb{P}(n)$ es un cono convexo abierto de $\mathcal{H}(n)$.

Proposición

Al conjunto $\mathbb{P}(n)$ se lo puede dotar de una estructura diferencial de modo que el espacio tangente en cada punto se puede identificar con $\mathcal{H}(n)$.

El espacio $\mathcal{H}(n)$ posee un producto escalar canónico dado por la traza

$$\langle A, B \rangle = \operatorname{tr}(AB),$$

el cual lo convierte en un espacio euclídeo. Por otro lado, recordemos que $\mathcal{G}l\,(n)$ actúa por conjugación en $\mathbb{P}(n)$, es decir, dada $G \in \mathcal{G}l\,(n)$

$$L_G(A) = G^* A G.$$

El conjunto $\mathbb{P}(n)$ es un cono convexo abierto de $\mathcal{H}(n)$.

Proposición

Al conjunto $\mathbb{P}(n)$ se lo puede dotar de una estructura diferencial de modo que el espacio tangente en cada punto se puede identificar con $\mathcal{H}(n)$.

El espacio $\mathcal{H}(n)$ posee un producto escalar canónico dado por la traza

$$\langle A, B \rangle = \operatorname{tr}(AB),$$

el cual lo convierte en un espacio euclídeo. Por otro lado, recordemos que $\mathcal{G}l\,(n)$ actúa por conjugación en $\mathbb{P}(n)$, es decir, dada $G \in \mathcal{G}l\,(n)$

$$L_G(A) = G^* A G.$$

Dado $A \in \mathbb{P}(n)$. En el tangente $T_A\mathbb{P}(n)$ definimos el producto escalar

$$\langle S, T \rangle_A = \operatorname{tr}\left(L_{A^{1/2}}^{-1}(S)\, L_{A^{1/2}}^{-1}(T)\right) = \left\langle L_{A^{1/2}}^{-1}(S), L_{A^{1/2}}^{-1}(T) \right\rangle.$$

En vez de $A^{1/2}$ se puede usar cualquier $G \in \mathcal{G}l\,(n)$ tal que $G^* G = A$.

Usando esta estructura métrica, si $\alpha : [a,b] \to \mathbb{P}(n)$ es una curva suave a trozos

$$\mathrm{Long}(\alpha) := \int_a^b \sqrt{\langle\, \dot{\alpha}(t), \dot{\alpha}(t)\,\rangle_{\alpha(t)}}\, dt$$

y dadas $A, B \in \mathbb{P}(n)$

$$\delta(A,B) = \inf\{\mathrm{Long}(\alpha) : \ \alpha \text{ es suave a trozos y une } A \text{ con } B\}.$$

Usando esta estructura métrica, si $\alpha : [a,b] \to \mathbb{P}(n)$ es una curva suave a trozos

$$\operatorname{Long}(\alpha) := \int_a^b \sqrt{\langle\, \dot{\alpha}(t), \dot{\alpha}(t)\,\rangle_{\alpha(t)}}\, dt$$

y dadas $A, B \in \mathbb{P}(n)$

$$\delta(A,B) = \inf\{\operatorname{Long}(\alpha) :\ \alpha \text{ es suave a trozos y une } A \text{ con } B\}.$$

Teorema

Sean $A, B \in \mathbb{P}(n)$. La geodésica que las une es la curva $\gamma : [0,1] \to \mathbb{P}(n)$ dada por

$$\gamma(t) = A^{1/2}(A^{-1/2} B\, A^{-1/2})^t A^{1/2}.$$

Usando esta estructura métrica, si $\alpha : [a,b] \to \mathbb{P}(n)$ es una curva suave a trozos

$$\mathrm{Long}(\alpha) := \int_a^b \sqrt{\langle\, \dot{\alpha}(t), \dot{\alpha}(t)\, \rangle_{\alpha(t)}}\, dt$$

y dadas $A, B \in \mathbb{P}(n)$

$$\delta(A,B) = \inf\{\mathrm{Long}(\alpha) :\ \alpha \text{ es suave a trozos y une } A \text{ con } B\}.$$

Teorema

Sean $A, B \in \mathbb{P}(n)$. La geodésica que las une es la curva $\gamma : [0,1] \to \mathbb{P}(n)$ dada por

$$\gamma(t) = A^{1/2}(A^{-1/2} B\, A^{-1/2})^t A^{1/2} = A \#_t B.$$

Usando esta estructura métrica, si $\alpha : [a, b] \to \mathbb{P}(n)$ es una curva suave a trozos

$$\mathrm{Long}(\alpha) := \int_a^b \sqrt{\langle \dot{\alpha}(t), \dot{\alpha}(t) \rangle_{\alpha(t)}}\, dt$$

y dadas $A, B \in \mathbb{P}(n)$

$$\delta(A, B) = \inf\{\mathrm{Long}(\alpha) : \ \alpha \text{ es suave a trozos y une } A \text{ con } B\}.$$

Teorema

Sean $A, B \in \mathbb{P}(n)$. La geodésica que las une es la curva $\gamma : [0, 1] \to \mathbb{P}(n)$ dada por

$$\gamma(t) = A^{1/2}(A^{-1/2} B\, A^{-1/2})^t A^{1/2} = A \#_t B.$$

A partir de esto se obtiene la siguiente caracterización geométrica de la media $A \# B$

Teorema

$$A \# B = \operatorname*{argmin}_{C \in \mathbb{P}(n)} \ \frac{\delta^2(C, A) + \delta^2(C, B)}{2}.$$

Sea (Ω, μ) un espacio de probabilidad y $1 \leq p < \infty$.

Definición

Decimos que $F : \Omega \to \mathbb{P}(n)$ es p-integrable si es medible y existe $A \in \mathbb{P}(n)$

$$\int_\Omega \delta^p(F(\omega), A)\, d\mu(\omega) < \infty.$$

Definimos los espacios $L^p(\Omega, \mathbb{P}(n))$ identificando funciones que coinciden μ-ctp.

Sea (Ω, μ) un espacio de probabilidad y $1 \leq p < \infty$.

Definición

Decimos que $F : \Omega \to \mathbb{P}(n)$ es p-integrable si es medible y existe $A \in \mathbb{P}(n)$

$$\int_\Omega \delta^p(F(\omega), A)\, d\mu(\omega) < \infty.$$

Definimos los espacios $L^p(\Omega, \mathbb{P}(n))$ identificando funciones que coinciden μ-ctp.

Dada $F \in L^2(\Omega, \mathbb{P}(n))$, definimos el baricentro de F del siguiente modo:

$$\beta_F = \underset{C \in \mathbb{P}(n)}{\operatorname{argmin}} \int_\Omega \delta^2(F(\omega), C)\, d\mu(\omega).$$

Si F adopta sólo dos valores A y B con igual probabilidad, entonces $\beta_F = A \# B$. Sin embargo, si toma tres o más valores no hay una fórmula cerrada para β_F.

Sea (Ω, μ) un espacio de probabilidad y $1 \leq p < \infty$.

Definición

Decimos que $F : \Omega \to \mathbb{P}(n)$ es p-integrable si es medible y existe $A \in \mathbb{P}(n)$

$$\int_\Omega \delta^p(F(\omega), A)\, d\mu(\omega) < \infty.$$

Definimos los espacios $L^p(\Omega, \mathbb{P}(n))$ identificando funciones que coinciden μ-ctp.

Dada $F \in L^2(\Omega, \mathbb{P}(n))$, definimos el baricentro de F del siguiente modo:

$$\beta_F = \underset{C \in \mathbb{P}(n)}{\operatorname{argmin}} \int_\Omega \delta^2(F(\omega), C)\, d\mu(\omega).$$

Si F adopta sólo dos valores A y B con igual probabilidad, entonces $\beta_F = A \# B$. Sin embargo, si toma tres o más valores no hay una fórmula cerrada para β_F.

Pregunta

Dadas $F, G \in L^2(\Omega, \mathbb{P}(n))$ tales que $F(\omega) \leq G(\omega)$ para μ-ctp, ¿vale que $\beta_F \leq \beta_G$?

Definición

Sea $F \in L^1(\Omega, \mathbb{P}(n))$. Dada $B \in \mathbb{P}(n)$ definimos

$$\beta_F := \underset{C \in \mathbb{P}(n)}{\operatorname{argmin}} \int_\Omega \delta^2(C, F(\omega)) - \delta^2(B, F(\omega))\, d\mu(\omega).$$

Definición

Sea $F \in L^1(\Omega, \mathbb{P}(n))$. Dada $B \in \mathbb{P}(n)$ definimos

$$\beta_F := \operatorname*{argmin}_{C \in \mathbb{P}(n)} \int_\Omega \delta^2(C, F(\omega)) - \delta^2(B, F(\omega))\, d\mu(\omega).$$

Comentarios: Sea Φ_B el funcional de la derecha, cuyo mínimo es β_F.

- El funcional Φ_B es continuo y si $\gamma : [0,1] \to \mathbb{P}(n)$ es una geodésica entonces

$$\Phi_B(\gamma(t)) \leq (1-t)\Phi_B(\gamma(0)) + t\Phi_B(\gamma(1)).$$

  Estas dos propiedades aseguran la existencia y unicidad del mínimo.
- La definición no depende de la B elegida. En efecto, notar que

$$\Phi_B(C) - \Phi_{B'}(C) = \int_\Omega \delta^2(B', F(\omega)) - \delta^2(B, F(\omega))\, d\mu(\omega)$$

  resulta independiente de C.
- Si $F \in L^2(\Omega, \mathbb{P}(n))$, entonces la definición coincide con la anterior.

Sean $\{X_k\}_{k\in\mathbb{N}}$ variables aleatorias i.i.d. y tal que $\mathbb{E}(|X_1|) < \infty$. Entonces

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow[n\to\infty]{} \mathbb{E}(X_1) \quad c.s.$$

Sean $\{X_k\}_{k\in\mathbb{N}}$ variables aleatorias i.i.d. y tal que $\mathbb{E}(|X_1|) < \infty$. Entonces

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow[n\to\infty]{} \mathbb{E}(X_1) \quad c.s.$$

Notemos que

$$S_3 = \frac{X_1 + X_2 + X_3}{3} = \frac{2}{3}\left(\frac{X_1 + X_2}{2}\right) + \frac{1}{3}X_3$$

$$S_4 = \frac{X_1 + X_2 + X_3 + X_4}{4} = \frac{3}{4}\left(\frac{X_1 + X_2 + X_3}{3}\right) + \frac{1}{4}X_4.$$

Ley fuerte de los grandes números
Caso escalar

Sean $\{X_k\}_{k\in\mathbb{N}}$ variables aleatorias i.i.d. y tal que $\mathbb{E}(|X_1|) < \infty$. Entonces

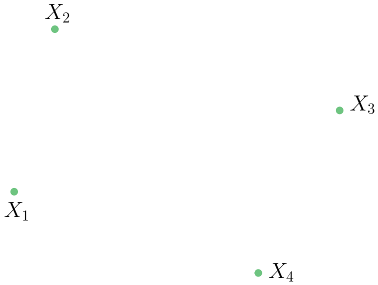
$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow[n\to\infty]{} \mathbb{E}(X_1) \quad c.s.$$

Notemos que

$$S_3 = \frac{X_1 + X_2 + X_3}{3} = \frac{2}{3}\left(\frac{X_1 + X_2}{2}\right) + \frac{1}{3} X_3$$

$$S_4 = \frac{X_1 + X_2 + X_3 + X_4}{4} = \frac{3}{4}\left(\frac{X_1 + X_2 + X_3}{3}\right) + \frac{1}{4} X_4.$$

Geométricamente



Sean $\{X_k\}_{k\in\mathbb{N}}$ variables aleatorias i.i.d. y tal que $\mathbb{E}(|X_1|)<\infty$. Entonces

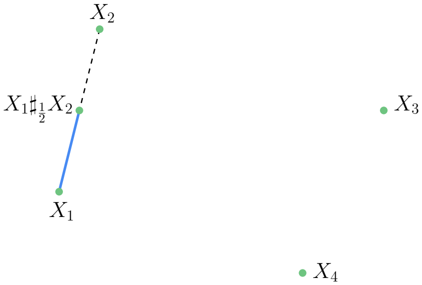
$$\frac{X_1+\ldots+X_n}{n}\xrightarrow[n\to\infty]{}\mathbb{E}(X_1)\quad c.s.$$

Notemos que

$$S_3=\frac{X_1+X_2+X_3}{3}=\frac{2}{3}\left(\frac{X_1+X_2}{2}\right)+\frac{1}{3}X_3$$

$$S_4=\frac{X_1+X_2+X_3+X_4}{4}=\frac{3}{4}\left(\frac{X_1+X_2+X_3}{3}\right)+\frac{1}{4}X_4.$$

Geométricamente



Sean $\{X_k\}_{k\in\mathbb{N}}$ variables aleatorias i.i.d. y tal que $\mathbb{E}(|X_1|) < \infty$. Entonces

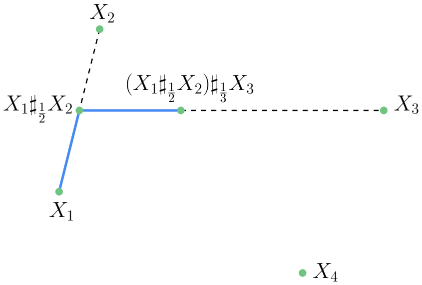
$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow[n\to\infty]{} \mathbb{E}(X_1) \quad c.s.$$

Notemos que

$$S_3 = \frac{X_1 + X_2 + X_3}{3} = \frac{2}{3}\left(\frac{X_1 + X_2}{2}\right) + \frac{1}{3}X_3$$

$$S_4 = \frac{X_1 + X_2 + X_3 + X_4}{4} = \frac{3}{4}\left(\frac{X_1 + X_2 + X_3}{3}\right) + \frac{1}{4}X_4.$$

Geométricamente



Sean $\{X_k\}_{k\in\mathbb{N}}$ variables aleatorias i.i.d. y tal que $\mathbb{E}(|X_1|) < \infty$. Entonces

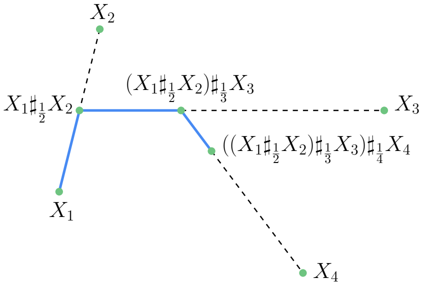
$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow[n\to\infty]{} \mathbb{E}(X_1) \quad c.s.$$

Notemos que

$$S_3 = \frac{X_1 + X_2 + X_3}{3} = \frac{2}{3}\left(\frac{X_1 + X_2}{2}\right) + \frac{1}{3}X_3$$

$$S_4 = \frac{X_1 + X_2 + X_3 + X_4}{4} = \frac{3}{4}\left(\frac{X_1 + X_2 + X_3}{3}\right) + \frac{1}{4}X_4.$$

Geométricamente



Definición

Sea $\{A_n\}_{n\in\mathbb{N}}$ una sucesión en $\mathbb{P}(n)$. Entonces definimos

$$
\begin{aligned}
S_1(A) &= A_1 \\
S_n(A) &= S_{n-1}(A)\#_{\frac{1}{n}}A_n \qquad (n \geq 2).
\end{aligned}
$$

Definición

Sea $\{A_n\}_{n\in\mathbb{N}}$ una sucesión en $\mathbb{P}(n)$. Entonces definimos

$$S_1(A) = A_1$$
$$S_n(A) = S_{n-1}(A)\#_{\frac{1}{n}} A_n \qquad (n \geq 2).$$

Teorema (Sturm '02)

Sean $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de vectores aleatorios i.i.d en $L^1(\Omega, \mathbb{P}(n))$. Entonces

$$S_n(X) \xrightarrow[n\to\infty]{} \beta_X \quad c.s.$$

Definición

Sea $\{A_n\}_{n\in\mathbb{N}}$ una sucesión en $\mathbb{P}(n)$. Entonces definimos

$$\begin{aligned} S_1(A) &= A_1 \\ S_n(A) &= S_{n-1}(A)\#_{\frac{1}{n}}A_n \qquad (n \geq 2). \end{aligned}$$

Teorema (Sturm ’02)

Sean $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de vectores aleatorios i.i.d en $L^1(\Omega, \mathbb{P}(n))$. Entonces

$$S_n(X) \xrightarrow[n\to\infty]{} \beta_X \quad c.s.$$

Corolario (Sturm ’02- Lawson, Lim ’11)

Dadas $F, G \in L^1(\Omega, \mathbb{P}(n))$ tales que $F(\omega) \leq G(\omega)$ para todo $\omega \in \Omega$. Entonces

$$\beta_F \leq \beta_G.$$

Dadas $A_1, \ldots, A_n \in \mathbb{P}(n)$, definamos

$$A = (A_1, \ldots, A_n, A_1, \ldots, A_n, A_1, \ldots, A_n, \ldots),$$

Teorema (Holbrook '12)

Si $F : \mathbb{Z}_n \to \mathbb{P}(n)$ *es la función definida como* $F(\bar{k}) = A_k$*, entonces*

$$\lim_{n\to\infty} S_n(A) = \beta_F.$$

Dadas $A_1, \ldots, A_n \in \mathbb{P}(n)$, definamos

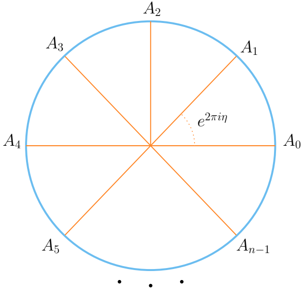
$$A = (A_1, \ldots, A_n, A_1, \ldots, A_n, A_1, \ldots, A_n, \ldots),$$

Teorema (Holbrook '12)

Si $F : \mathbb{Z}_n \to \mathbb{P}(n)$ es la función definida como $F(\bar{k}) = A_k$, entonces

$$\lim_{n\to\infty} S_n(A) = \beta_F.$$

- Identificando $\mathbb{Z}_n$ con las raíces n-ésimas de la unidad, este caso periódico sugiere estudiar las rotaciones irracionales
- Pensando $\bar{k} \mapsto \overline{k+1}$ como una transformación ergódica en $\mathbb{Z}_n$, el resultado de Holbrook es un teorema ergódico.



Teoremas ergódicos en términos de medias inductivas

Sea $(G,+)$ un grupo abeliano compacto y $h \in G$ de modo que $\tau(g) = g + h$ resulta ergódica en G.

Dada $A : G \to \mathbb{P}(n)$ definimos $\mathcal{O}_A : G \to \mathbb{P}(n)^{\mathbb{N}}$ del siguiente modo:

$$\mathcal{O}_A(g) := \big(A(g), A(\tau(g)), A(\tau^2(g)), A(\tau^3(g)), \ldots\big).$$

Teoremas ergódicos en términos de medias inductivas

Sea $(G,+)$ un grupo abeliano compacto y $h \in G$ de modo que $\tau(g) = g + h$ resulta ergódica en G.

Dada $A : G \to \mathbb{P}(n)$ definimos $\mathcal{O}_A : G \to \mathbb{P}(n)^{\mathbb{N}}$ del siguiente modo:

$$\mathcal{O}_A(g) := \big(A(g), A(\tau(g)), A(\tau^2(g)), A(\tau^3(g)), \dots\big).$$

Teorema (A.-Ghiglioni-Stojanoff)

Dada $A \in L^1(G, \mathbb{P}(n))$, para casi todo $g \in G$

$$\lim_{n\to\infty} S_n(\mathcal{O}_A(g)) = \beta_A.$$

Teoremas ergódicos en términos de medias inductivas

Sea $(G, +)$ un grupo abeliano compacto y $h \in G$ de modo que $\tau(g) = g + h$ resulta ergódica en G.

Dada $A : G \to \mathbb{P}(n)$ definimos $\mathcal{O}_A : G \to \mathbb{P}(n)^{\mathbb{N}}$ del siguiente modo:

$$\mathcal{O}_A(g) := \big(A(g), A(\tau(g)), A(\tau^2(g)), A(\tau^3(g)), \dots\big).$$

Teorema (A.-Ghiglioni-Stojanoff)

Dada $A \in L^1(G, \mathbb{P}(n))$, para casi todo $g \in G$

$$\lim_{n\to\infty} S_n(\mathcal{O}_A(g)) = \beta_A.$$

Teorema (A.-Ghiglioni-Stojanoff)

Dado $A \in L^p(G, \mathbb{P}(n))$

$$\lim_{n\to\infty} \int_{\mathbb{T}} \delta^p(S_n(\mathcal{O}_A(g)), \beta_A)\, dg = 0.$$

Fin